A New, Modular Mass Calibrant for High-Mass MALDI-MS

Weidmann, Simon ; Barylyuk, Konstantin ; Nespovitaya, Nadezhda ; Mädler, Stefanie ; Zenobi, Renato

Abstract: The application of matrix-assisted laser desorption/ionization mass spectrometry (MALDIMS) for the analysis of high-mass proteins requires suitable calibration standards at high m/z ratios. Several possible candidates were investigated, and concatenated polyproteins based on recombinantly expressed maltodextrin-binding protein (MBP) are shown here to be well suited for this purpose. Introduction of two specific recognition sites into the primary sequence of the polyprotein allows for the selective cleavage of MBP3 into MBP and MBP2. Moreover, these MBP2 and MBP3 oligomers can be dimerized specifically, such that generation of MPB4 and MBP6 is possible as well. With the set of calibrants presented here, the m/z range of 40–400 kDa is covered. Since all calibrants consist of the same species and differ only in mass, the ionization efficiency is expected to be similar. However, equimolar mixtures of these proteins did not yield equal signal intensities on a detector specifically designed for detecting high-mass molecules.

DOI: https://doi.org/10.1021/ac400129h

A New, Modular Mass Calibrant for High-Mass MALDI-MS

Simon Weidmann,[†] Konstantin Barylyuk,[†] Nadezhda Nespovitaya,[†] Stefanie Mädler,[†,‡] and Renato Zenobi[*,†]

Department of Chemistry and Applied Biosciences, ETH Zurich, CH-8093 Zurich, Switzerland

E-mail: zenobi@org.chem.ethz.ch

Abstract

The application of matrix-assisted laser desorption/ionization mass spectrometry (MALDI-MS) for the analysis of high-mass proteins requires suitable calibration standards at high *m/z* ratios. Several possible candidates were investigated, and concatenated polyproteins based on recombinantly expressed maltodextrin-binding protein (MBP) are shown here to be well suited for this purpose. Introduction of two specific recognition sites into the primary sequence of the polyprotein allows for the selective cleavage of MBP_3 into MBP and MBP_2. Moreover, these MBP_2 and MBP_3 oligomers can be dimerized specifically, such that generation of MPB_4 and MBP_6 is possible as well. With the set of calibrants presented here, the *m/z* range of 40–400 kDa is covered. Since all calibrants consist of the same species and differ only in mass, the ionization efficiency is expected to be similar. However, equimolar mixtures of these proteins did not yield equal signal intensities on a detector specifically designed for detecting high-mass molecules.

[*]To whom correspondence should be addressed
[†]Department of Chemistry and Applied Biosciences, ETH Zurich, CH-8093 Zurich, Switzerland
[‡]Current address: Centre for Research in Mass Spectrometry, York University, Toronto, ON M3J 2R7, Canada

Introduction

Mass Spectrometry (MS) has emerged as a very powerful analytical technique for studying large proteins or protein complexes.[1,2] In order to keep proteins intact, soft ionization techniques such as electrospray ionization (ESI) or matrix-assisted laser desorption/ionization (MALDI) are widely applied. In contrast to ESI-MS where highly charged ions are formed, MALDI forms predominantly singly charged ions. The *m/z* ratio of the analyte therefore becomes rather high and requires the use of special high-mass detectors that are capable of detecting very large ions.[3,4]

Mass spectrometers are usually calibrated using samples with properties similar to the analyte and well-known masses as standards. When peptides are measured, calibration standards for a broad *m/z* range are available, and high mass accuracy is achievable.[5,6] However, when high-mass proteins are investigated, and special high-mass detectors that extend the *m/z* range to approximately 1.5 MDa are used, the choice of calibrants is much more limited. Bovine serum albumin,[4,7] immunoglobulin G,[4] or a commercial high molecular weight mass calibrant kit (Invitromass)[8,9] can be employed, but in many situations, the *m/z* range covered is too narrow; new calibrants are therefore necessary. Some recently developed new calibrants are either based on recombinant van Willebrand factor, a protein with a molecular weight of about 260 kDa,[10] or nonspecific aggregation of medium size proteins.[11] Furthermore, a modular mass calibrant based on recombinant protein building blocks without PTMs has been developed by Agilent Technologies for their automated lab-on-a-chip platform.[12,13] Although this calibrant might be suitable for MALDI-MS it has never been commercialized as such.

An accurate and reliable calibration is useful for a broad range of applications, first and foremost for exact mass determination. Since MALDI-MS is a very fast and accurate method for mass determination and only minute amounts of sample are needed, it has become the method of choice for quality control of expressed proteins. In comparison to other analytical techniques such as SDS-PAGE, capillary gel electrophoresis, or size exclusion chromatography, MALDI MS provides superior mass resolution and accuracy independent of the three-dimensional structure of the analyte.[14–16] High accuracy mass information is key, for example, for determining the stoi-

chiometry of protein–protein,[3,4,17,18] protein–DNA,[19] or protein–RNA complexes.[20] Exact mass information is also used for the determination of the degree and the kind of post-translational modifications (PTMs),[7] PEGylations,[21] as well as for determining truncation of the primary sequence. Selective truncation is often used in protein crystallography in order to reduce the floppy ends of a protein and therefore facilitate the crystallization of proteins.[22–25] With MALDI-MS the extent and in some cases even the location of truncation can be determined by measurement of the molecular weight and comparison with the primary sequence, provided that high enough mass accuracy is achieved.

A good calibration standard for the high-mass range must fulfill several requirements: *i)* it should have a high molecular weight, *ii)* it should be detected with good sensitivity, *iii)* its signals should be well resolved; *iv)* the calibrant should be easy to use, *v)* monodisperse, and *vi)* of high purity. Potential candidates that fulfill the first two requirements are either monomeric proteins with a high mass or multimeric proteins with an intermediate mass per subunit. Since noncovalent complexes are easily disrupted in MALDI, either nonspecific multimerization or chemical stabilization with crosslinkers is necessary, which, however, broadens the signals. Besides the peak broadening due to nonspecific aggregation this approach has another disadvantage: the aggregate is formed during MALDI from individual subunits. Therefore, the conditions are different than for the formation of monomeric ions and some small errors might be introduced. However, for daily use this method might still be applicable.

One of the problems that arises when using high-mass proteins is the width of the signal. Often, those proteins are not recombinantly expressed, and generally carry PTMs. These PTMs are heterogeneous, which leads to substantial peak broadening.

Our goal was do design a modular system for calibration. This has the advantage that not every calibrant needs to be expressed separately, but can be synthesized from the expressed starting protein in only a few steps (see Figure 1). To design such a modular system, a specific repeat unit was selected. Several copies of this repeat unit can in this case be cloned into the same expression vector and expressed in suitable cells.[26] For our experiments the protein chosen as repeating unit

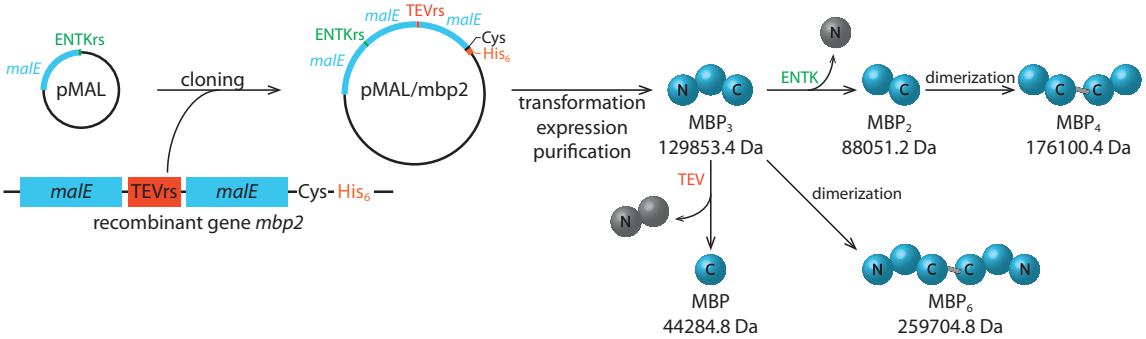


Figure 1: Schematic representation of the newly designed concatenated polyprotein based on maltodextrin-binding protein (MBP). The introduction of two recognition sites for different proteases (ENTKrs, TEVrs) allows for the orthogonal cleavage of the trimeric species. Due to the cysteine residue that is present, a selective dimerization of MBP_2 and MBP_3 into MBP_4 and MBP_6, respectively, is possible. The molecular weights of the different MBP species have been calculated based on the primary sequence using the ExPASy Proteomics Tool.

was maltodextrin-binding protein (MBP), a globular 40 kDa protein[27] that is routinely used as a fusion partner in protein expression.[28] This protein is very stable in solution and can easily be submitted to affinity-based purification on amylose resin. A recombinant gene *mbp2* was inserted into a commercially available pMAL expression vector downstream of the MBP encoding *malE* gene and the following enterokinase recognition site (ENTKrs). The gene *mbp2* consists of two copies of *malE*, which are separated by a tobacco edge virus protease recognition site (TEVrs). Following the two MBP encoding genes, one single cysteine residue was inserted into *mbp2*. Since this is the only cysteine in the whole protein sequence, it can be used for selective dimerization of MBP_3 via a disulfide bridge. On the C-terminus of the genetically engineered polyprotein, a polyhistidine tag (His_6, His-tag) was introduced to allow for purification of the targeted protein using nickel affinity chromatography. The whole primary sequence of the concatenated polyprotein is shown in the Supporting Information (Figure S1).

In a recent publication, Phillips *et al.* showed the possibility to produce fibrous nanostructures by reaction of recombinantly expressed protein building blocks into longer assemblies. The tetratricopeptide repeats used self-assemble into stable superhelices. The length (and thus mass) is controlled by the number of repeat motifs used.[29] A whole library of repeat proteins has been developed by Plückthun and coworkers. Their designed ankyrin repeat proteins (DARPins) con-

sist of ankyrin repeat modules with different numbers of repeats and end caps on the C- and the N-terminus.[30,31]

The use of concatenated proteins is also useful in other fields. When including several copies of the domain of interest into a single protein, insights into the chemical and mechanical stability of proteins are gained and finally lead to information about the structure of the protein of interest.[32–37]

In this work several possible calibration standards were investigated and crucial issues for successful calibration were studied. Since no commercially available protein suited the requirements mentioned above, the new modular mass calibrant based on recombinantly expressed polyprotein was developed. Expression of a genetically engineered pMAL/mpb2 vector in *Escherichia coli* gave a covalently bound trimer of MBP with specific recognition sites for proteolytic enzymes and dimerization sites allowing for easy synthesis of calibrants with different masses from one single protein. The ease of use of this modular system was found to have great advantages over other calibrants. Commercialization of this product can simplify calibration of high-mass MALDI-MS, and analysis of heavy intact proteins could become a routine task.

Experimental Section

Materials

Sinapinic acid (SA), acetonitrile (ACN), DL-dithiothreitol (DTT), imidazole, glycerol, lysozyme and phosphate buffered saline tablets (pH 7.4) were bought from Sigma-Aldrich (Buchs, Switzerland), trifluoroacetic acid (TFA) from Acros Organics (Geel, Belgium), and sodium chloride (NaCl) from J. T. Baker (Deventer, Netherlands). Amicon Ultra-0.5 centrifugal filters (NMWL 30k and 50k) and ammonium acetate (NH_4OAc) were purchased from Merck (Darmstadt, Germany). All chemicals were of the highest purity available and were used without further purification. For nickel affinity purification on a small scale, a His SpinTrap Kit (GE Healthcare Life Sciences, Glattbrugg, Switzerland) was used. LB broth was purchased from BD (Franklin Lakes, NJ, USA), ampicillin, isopropyl-β-D-thiogalactopyranoside (IPTG), and phenylmethylsulfonyl fluo-

ride (PMSF) were of 'BioChemica' quality and obtained from AppliChem (Darmstadt, Germany). Bovine thyroglobulin (Tg bov) was purchased from Morphosys AbD (Düsseldorf, Germany), recombinant human fibronectin (rh Fn) from R&D Systems Europe (Abingdon, UK), and phosphorylase B (lot: 122110AW1) from Proteabio Europe (Langlade, France). GroEL was a gift from Philippe Ringler (University of Basel, Basel, Switzerland), protein histidine phosphatase (PHP) was purchased from the lab of Prof. Dr. Simone König (University of Münster, Germany). The plasmid pMAL/mbp2 for MBP_3 expression was manufactured by GenScript Corp. (Piscataway, NJ, USA). The enzymes AcTEV Protease and enterokinase, light chain (ENTK) were obtained from Invitrogen (Lucerne, Switzerland) and from BioConcept (Allschwil, Switzerland), respectively. The expression was performed in chemically competent BL21 Star *E. coli* cells from Invitrogen (Lucerne, Switzerland). Nanopure water (18.2 MΩ cm) was prepared using a NANOpure Diamond water purification system (Barnstead International, Dubuque, IA, USA).

Expression and Purification of Concatenated Polyproteins

50 µL of *E. coli* cell suspension was mixed with 1 µL of plasmid. After incubation for 25 min on ice and subsequent heat shock (30 s, 42 °C), 200 µL of LB broth were added and the mixture was incubated on a shaker at 37 °C for 1 h. An aliquot of 50 µL was distributed on an agar plate and incubated for 16 h at 37 °C. To pre-cultivate the cells, a single colony was transferred into 5 mL LB broth containing 500 µg ampicillin. After incubation of the preculture for 16 h at 37 °C on a shaker, an aliquot of 800 µL was mixed with 200 µL of 30 % glycerol and stored at −80 °C for later expression. The remaining pre-culture mixture was transferred into 1 L of LB broth and incubated at 37 °C. Cell growth was monitored by measuring the optical density at 600 nm (OD_{600}). Once OD_{600} was > 0.8, protein expression was started by addition of 1 mL of 1 M IPTG.

After 4 h of expression, the cell suspension was centrifuged and the cell pellets were resuspended in 25 mL of binding buffer (20 mM phosphate buffered saline buffer (PBS), 100 mM NaCl, 20 mM imidazole, 1 mM DTT). PMSF was added to this suspension a final concentration of 1 mM as well as 20 µL of 0.5 mg/mL lysozyme for cell lysis. The suspension was homogenized in a

microfluidizer and subsequently centrifuged for 30 min with 14 000 g at 4 °C.

The supernatant obtained by centrifugation was divided into three parts, which were purified separately using nickel affinity chromatography (ÄKTAprime equipped with 5 mL HisTrap FF, GE Healthcare Life Sciences, Glattbrugg, Switzerland). The system was equilibrated before each purification run with the binding buffer mentioned above. Approximately 10 mL of supernatant was loaded to the column. After elution of proteins without His-tag, the desired proteins were eluted by switching to an elution buffer (20 mM PBS, 100 mM NaCl, 300 mM imidazole, 1 mM DTT). To desalt the eluted protein, a desalting column (HiTrap Desalting, GE Healthcare Life Sciences, Glattbrugg, Switzerland) was used, the protein was injected into the system, and eluted using desalting buffer (20 mM PBS, 100 mM NaCl, 1 mM DTT). The flow rate was kept constant at 10 mL/min. The final protein concentration was determined as 2 mg/mL by measuring the UV absorption at 280 nm. For storage, the protein solution was filtered using a 2 µm filter, split into several aliquots of 1 mL, and freeze-dried using a lyophilizer (Alpha 2-4 LD plus, Christ, Osterode am Harz, Germany). To avoid problems with purification later on, no protection agents such as glycerol or polyethylene glycol were added.

After purification, the proteins are stable in solution for several months when stored at 4 °C. However, the cysteine moiety introduced for dimerization is prone to oxidation. Therefore, dimerization reactions should be performed directly after expression and/or purification. Dimerized species (MBP_4 and MBP_6) are less soluble than the expressed MBP_3. Thus, precipitation can occur and the proteins should be used directly after reaction and cannot be stored for longer time. When reconstituting MBP_3 from lyophilized powder, either pure water or 1 mM DTT can be used. In either case, it is important not to shake or vortex the mixture to prevent precipitation of the reconstituted protein.

Sample Preparation

Tg bov, GroEL, and PHP were used as delivered. Lyophilized rh Fn and phosphorylase B were reconstituted at concentrations of 100 µg/mL in PBS and of 0.5 mg/mL in water, respectively. All

samples were diluted with water or PBS to the desired concentrations. MBP_3 was either used directly, or enzymatically cleaved into MBP_2 and MBP. This was achieved by incubation of 10 nmol MBP_3, either with 1 µL AcTEV protease (10 U/mL) or with 6 µL of 2 µg/mL enterokinase for 16–24 h at 25 °C. The MBP_3 concentration used was typically 1–5 µM if not noted otherwise.

To prepare the matrix solution, SA was dissolved at a concentration of 10 mg/mL in acetonitrile/water/TFA (49.5:49.5:1, v:v:v). For MALDI analysis, typically 1 µL of the sample under investigation was carefully mixed with 1 µL of matrix solution. 0.5–1 µL of this mixture was deposited on a stainless steel MALDI plate and allowed to crystallize.

MALDI-MS

For all experiments, a MALDI TOF/TOF mass spectrometer (4800 Plus, AB SCIEX, Darmstadt, Germany) equipped with a commercially available high-mass detector (HM2tuvo, CovalX, Zürich, Switzerland) was used. The system was operated in linear mode, the effective flight path including the high-mass detector was 1.7 m. The default delayed extraction settings that were provided by the manufacturer (2000 ns) were used. This mass spectrometer is equipped with a frequency-tripled Nd:YAG laser (355 nm) and the laser power was adjusted to lie just above the threshold for ion formation. To obtain a mass spectrum, 100–500 laser shots, recorded automatically at random spot positions, were accumulated.

Results and Discussion

Protein Expression

The expressed protein was harvested and characterized by MALDI-MS. The expression was found to yield a mixture of MBP and MBP_3 (Figure 2a), which was further purified using gel filtration chromatography (ÄKTAprime equipped with Superdex 200 10/300 GL, GE Healthcare Life Sciences, Glattbrugg, Switzerland). The fractions obtained contained either monomeric or trimeric

MBP of very high purity (see Figure 2). It was impossible to use nickel affinity chromatography to separate this mixture, since both species carry a His-tag on their C-terminus. This indicates that the MBP monomer is not produced due to incomplete protein synthesis, since in this case the C-terminal His-tag of the sequence would be missing and MBP would be removed during the purification. One possible reason that the expression yields a mixture of MBP and MBP_3 is probably due to the primary sequence of the polyprotein. In front of each MBP encoding domain, a methionine residue is situated. Since the expression vector is quite large, it might fold and the ribosome might "jump" from the starting sequence in front of the first MBP to the methionine residue in front of the last MBP, to express only a single MBP protein. The expression of MBP_2 is not observed because the expression vector is not able to fold in such a way that the above mentioned "jump" to the methionine in front of the second MBP encoding gene is possible.

To test for the orthogonal cleavability of MBP_3, the trimeric species was subjected to proteolytic digestion using ENTK (data not shown) and AcTEV. As can be seen in Figure 3c, the mixture obtained after treatment with AcTEV contained MBP and MBP_2. The C-terminal His-tag allows for purification using nickel affinity chromatography. In this case, the flow-through of the His SpinTrap was collected and concentrated (Figure 3d), and subsequent buffer exchange to 100 mM NH_4OAc buffer produced a series of specific and nonspecific multimers. The exchange of non-volatile PBS buffer by volatile NH_4OAc does not reduce the peak width, indicating that salt adducts are not the main reason for the observed peak width.

The buffer of purified MBP_3 was exchanged to DTT-free PBS buffer using Amicon centrifugal filters. The cysteine residues of MBP_3 formed a disulfide bond and yielded MPB_6 (Figure 3b). The hexameric MBP species is prone to precipitation and can therefore not be stored in solution for a long time.

Comparison of Different Calibrants

Evaluation of other potential candidates for calibration standards showed that several problems exist. When using multimeric proteins with high mass such as Tg bov, which is a homodimer with

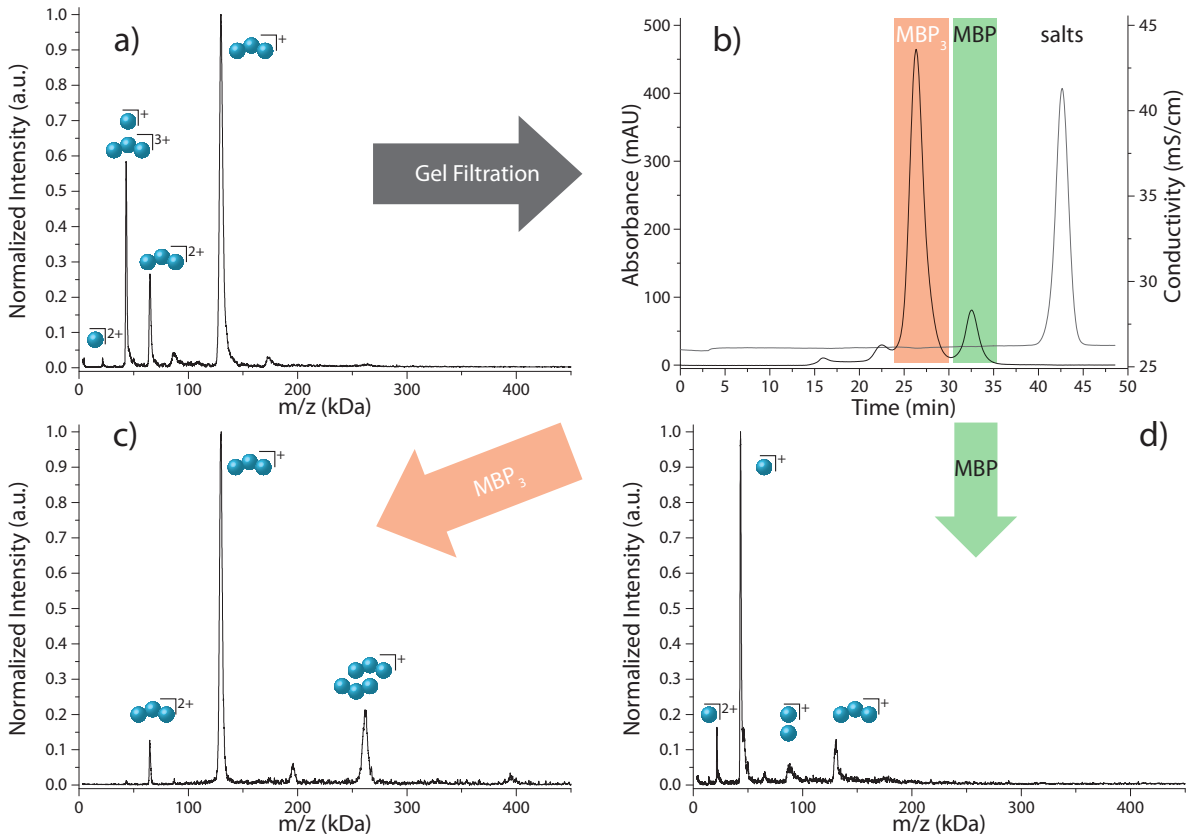


Figure 2: Expression of the pMAL/mpb2 vector in *E. coli* yielded a mixture of MBP and MBP_3 (subfigure a). Purification with gel filtration chromatography (b) allowed separation of four different peaks. The first and the second peaks are probably due to nonspecific multimerization and dimerization, respectively. However, due to low concentrations of these fractions no MALDI mass spectra could be recorded. The fractions eluting around 27 min and 32 min could be identified using MALDI-MS as MBP_3 (c) and MBP (d), respectively.

a total mass of 660 kDa, often only broad signals were observed. In case of the monomeric species, a peak width at half height (FWHM) of more than 20 kDa was measured (Figure 4). Although a dimer of Tg bov was observed as well, the signal was too wide to be useful as calibrant. One of the heaviest recombinantly expressed proteins that is commercially available is rh Fn with a molecular weight of 220 kDa. The FWHM of its molecular ion peak is about half that of Tg bov, a bit more than 11 kDa, which is a reasonably narrow for a glycoprotein, but is still much too wide to be useful as calibrant. Another drawback of rh Fn is the fact that no nonspecific multimers could be observed, therefore only a *m/z* range up to 220 kDa would be covered.

Two different multimeric complexes, PHP and GroEL, with molecular weights of the monomers

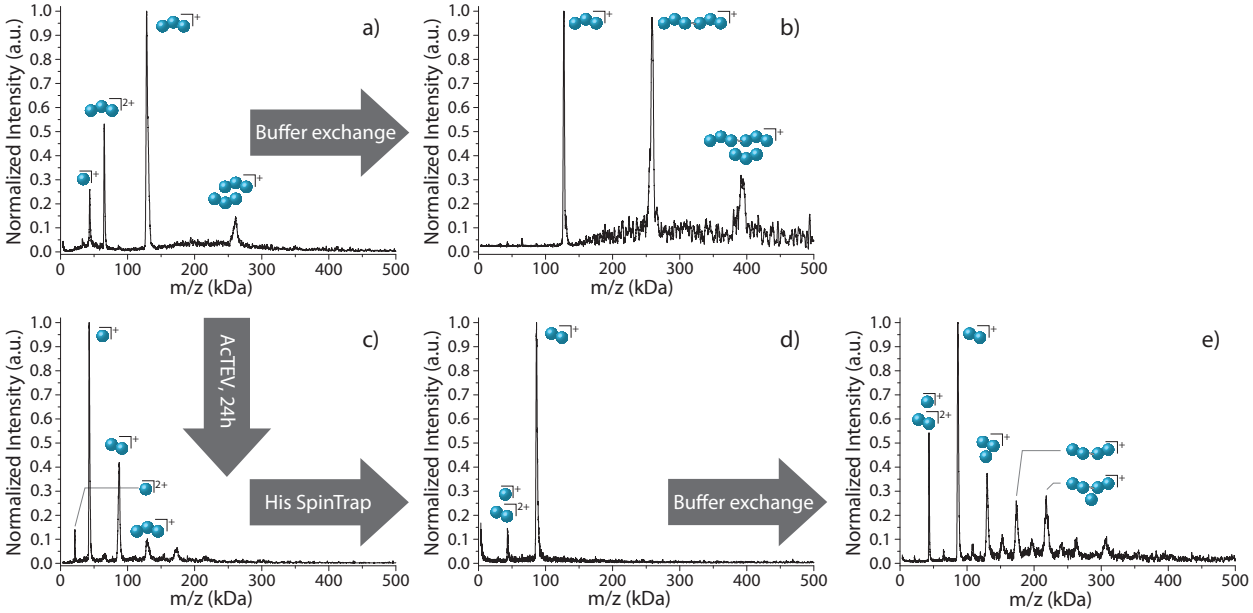


Figure 3: When analyzing MBP_3 with high-mass MALDI-MS, signals with a very narrow peak width of only 2.2 kDa are observed (a). Successful dimerization is shown in subspectra b) of this figure. Removing DTT using Amicon centrifugal filters leads to formation of MBP_6. This dimerization does not lead to a significant peak broadening, and due to the nonspecific clustering of MBP_6 the accessible *m/z* range of the calibration extends to 400 kDa. Subspectra c), d), and e) show the synthesis of MBP_4 including the intermediate steps after proteolytic digest using AcTEV (c), and purification using nickel affinity chromatography (d). When using AcTEV as protease, the His-tag remains on the MBP species. To obtain MBP_2 the flow-through was collected and concentrated. Buffer exchange to DTT-free NH_4OAc buffer allowed for the synthesis of specific and non-specific multimers, such as MBP_4.

of 16 and approximately 57 kDa, respectively, were investigated as well. Both proteins showed nonspecific clustering, i.e., a "calibration ladder" with well-resolved signals was observed. In the case of PHP, up to dodecamers could be measured. These findings are in good agreement with the work of Ludwig *et al.*[11] In their publication, they also found PHP aggregates with masses of more than 600 kDa when performing gel filtration chromatography, but they were not able to detect these aggregates with MS. Double-checking their findings using our MALDI-MS system gave the same results: no aggregates exceeding 200 kDa could be observed. Detection of the heavier aggregates probably was not successful either because the concentration was too low, or because the interactions forming multimers are mostly noncovalent and therefore disrupted during MALDI.

In the case of GroEL, clusters only up to the pentamer could be detected. Nevertheless, due

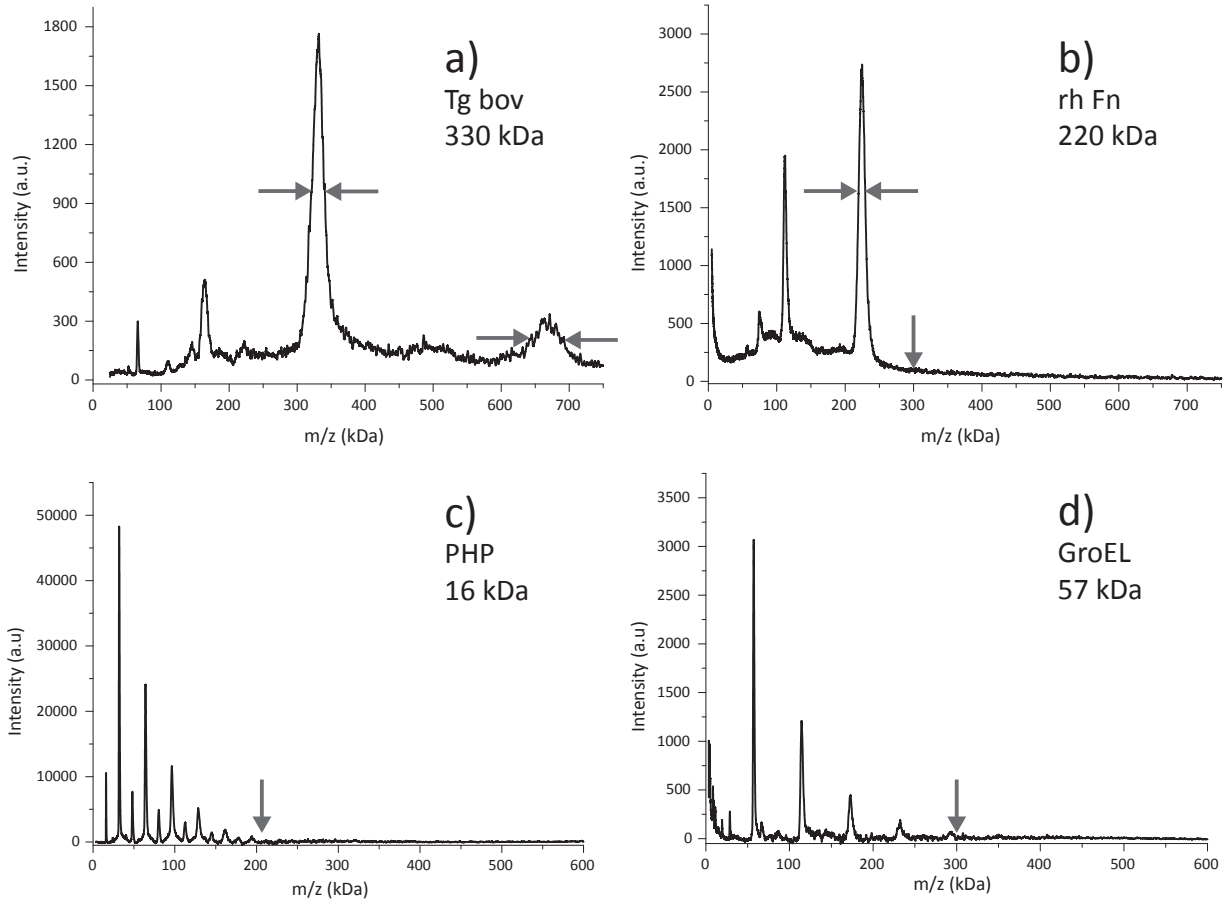


Figure 4: Comparison of different potential calibration standards: a) bovine thyroglobulin (MW: 330 kDa), b) recombinant human fibronectin (MW: 220 kDa), c) protein histidine phosphatase (MW: 16.6 kDa per monomer), d) GroEL (MW: 57 kDa per monomer). In the upper two cases the signals were too broad to be useful as calibrant (indicated by horizontal arrows). The vertical arrow indicates the end of the calibration region. For none of the proteins presented here it was possible to achieve a calibration with sufficiently well resolved signals and a calibration range up to approximately 400 kDa.

to the higher mass of the monomer, this leads to an upper end of the calibration of approximately 300 kDa.

Using the recombinantly expressed and purified MBP_3 it is possible to obtain spectra with well defined signals and a very narrow peak width of only 2.2 kDa for the trimer (see Table 1). Since nonspecific multimers are also observed, the *m/z* range that is covered can be extended up to 400 kDa.

There are many factors influencing the peak width in MALDI-MS.[38] An important, although

minor contribution comes from the isotopic distribution pattern (see Supplementary Information S2). Of much more influence are cation and matrix adducts that are formed during MALDI. For example, formation of a matrix adduct with SA adds 206 Da to the molecular weight.[39] Another aspect leading to a pronounced peak width when performing MALDI-MS in linear mode is the difference in kinetic energies after applying the acceleration pulse. This difference originates from different positions of the newly formed ions in the MALDI plume and leads to a spread in flight times and therefore peak broadening.

Table 1: Summary of the potential calibration candidates under investigation. The name, the theoretical and measured molecular weights (MW_{th} and MW_{meas}), the observed full width at half maximum (FWHM), and the calculated resolution R (MW_{meas}/FWHM) are listed.

Name	MW_{th}[a] (kDa)	MW_{meas} (kDa)	FWHM (kDa)	R
Bovine Thyroglobulin	303.2	331.8	21.5	15.4
Recombinant Human Fibronectin	206.6	224.8	11.4	19.7
GroEL – Trimer	172.0	172.5	4.6	37.5
GroEL – Tetramer	229.3	232.4	5.4	43.0
Maltodextrin-Binding Protein – Trimer (MBP_3)	129.9	128.3	2.2	58.3
Maltodextrin-Binding Protein – Hexamer (MBP_6)	259.8	258.5	4.3	60.1

[a] MW_{th} is calculated based on the primary sequence without any PTMs.

To investigate the relative response, purified MBP and MBP_3 were mixed in an equimolar ratio. Good signals could easily be detected using concentrations as low as 0.4 µM of total protein concentration (Figure 5 bottom). The dimer MBP_2 was omitted in this mixture to avoid superposition of the signals of MBP_2 and nonspecific MBP dimer. Although an equimolar mixture was investigated, the signal of MBP_3 shows higher signal intensity than the monomeric species. This could be either due to the fact that a trimeric species contains more basic sites and is therefore more easily protonated compared to the monomer. Another and more probable explanation is the increased sensitivity towards heavier ions of the detector used. From our experience, this detector works best for *m/z* ratios above 60 000. The CovalX HM2tuvo works as an ion-to-ion conversion dynode (ICD) where the impact of an ion generates secondary ions that are further accelerated towards an electron multiplier.[40] It is known that in cluster secondary ion mass spectrometry, the impact of

heavy cluster primary ions leads to an increased secondary ion yield.[41] In an ICD, an increased yield of secondary ions is also expected with increasing mass of the impinging primary ion, which will lead to a higher number of electrons and therefore to a higher signal intensity. However, the identification of factors influencing the relative response of protein in MALDI-MS is beyond the scope of this publication and will be investigated in the near future.

To test the suitability of internal calibration with the MBP/MPB_3 mixture (0.4 µM final concentration of each protein), phosphorylase B was added to a final concentration of 1 µM. Measurements using the default calibration of the instrument yielded molecular weights of approximately 130 kDa for MBP_3 and 97.6 kDa (mass error: 0.2 %) for phosphorylase B. The spectrum was recalibrated using the Data Explorer Software (Version 4.9, AB SCIEX, Darmstadt, Germany) using the signals of MBP and MBP_3 as internal calibrants. The mass of phosphorylase B was determined as 97 327 Da (mass error: 0.07 %) in this way, which matches the molecular weight (97.4 kDa) given by the manufacturer in the certificate of quality assurance for this lot better. The mass error of about 70 Da is lower than the molecular weight of almost all amino acids. The determination of molecular masses with such precision could allow drawing conclusions on the primary sequence of the investigated sample or, more specifically, determine truncations. Another application is in the quality control of therapeutic agents, where the degree of glycosylation of such agents could be controlled.

Depending on the requirements of the measurement, the calibrants presented here can be used either as internal or as external calibrants. When external calibration is applied, usually the calibrants are measured first and the correction factors are stored for subsequent measurements. With the different calibrants MBP, MBP_2, MBP_3, MBP_4, and MBP_6 different mixtures can be prepared according to the *m/z* range that is needed. The same holds true for internal calibration where it is important to avoid superposition of the calibrant and the analytes under investigation.

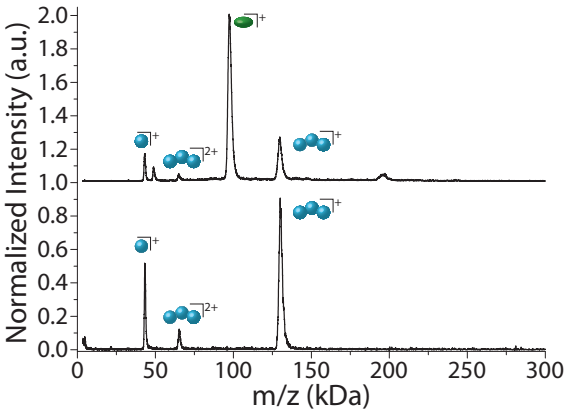


Figure 5: In the bottom spectrum a mixture of MBP and MBP_3 with final concentrations of 0.2 µM each was measured. Although the sample contained equal amounts of MBP and MBP_3, the signal of MBP_3 is more intense. This might be due to the increased sensitivity of the ICD detector to higher-mass primary ions: heavier ions lead to more secondary ions and therefore a more intense signal. To test for the feasibility of our modular calibration system for internal calibration a mixture containing MBP, MBP_3, and phosphorylase B (green ellipse) with concentrations of 0.4 µM, 0.4 µM, and 1 µM, respectively, was investigated. In the upper spectrum the corresponding spectrum is shown. The signals of MBP and MBP_3 were used for mass calibration to determine the mass of phosphorylase B.

Conclusions

Calibration of MALDI-MS in the high mass range is not straightforward. In the work presented here, different calibration standards were compared. It was shown that certain proteins are not suitable for calibration due to their wide signals caused by glycosylation (Tg bov), heterogeneities (rh Fn), or their limited mass range (rh Fn, PHP, GroEL). Two main conclusions are drawn from these observations: *i)* the protein should be recombinantly expressed and not carry any PTMs in order to yield narrow signals, and *ii)* nonspecific clustering leads to ladder-like signals over a wide *m/z* range, which is very useful for calibration in the *m/z* range above the MW of the molecular ion. Based on these findings a new modular calibrant was designed. Instead of expressing one huge protein, a concatenated polyprotein of medium sized proteins was used. The design was chosen in a way that the polyprotein could be orthogonally cleaved into subunits and either the whole polyprotein or its subunits could be dimerized into larger assemblies. Using this system, it was found to be possible to calibrate a mass range from 40 to approximately 400 kDa with

only one expressed protein and a maximum of two reaction steps. The FWHMs of the MBP protein oligomers presented here are in the range of 2–5 kDa, which is significantly lower than any other calibrant used before. Using these signals and the exactly known molecular mass of the investigated protein, it is possible to achieve a calibration over a wide *m/z* range with unprecedented accuracy and precision.

Acknowledgement

The authors thank the SystemX.ch initiative (CINA grant) and the Swiss National Science Foundation (grant no. 200020-124663) for their financial support.

References

(1) Bich, C.; Zenobi, R. *Curr. Opin. Struct. Biol* **2009**, *19*, 632–639.

(2) Mädler, S.; Boeri Erba, E.; Zenobi, R. *MALDI-ToF Mass Spectrometry for Studying Noncovalent Complexes of Biomolecules*; Topics in Current Chemistry; Springer Berlin/Heidelberg, 2012; pp 1–36.

(3) Nazabal, A.; Wenzel, R. J.; Zenobi, R. *Anal. Chem.* **2006**, *78*, 3562–3570.

(4) Pimenova, T.; Nazabal, A.; Roschitzki, B.; Seebacher, J.; Rinner, O.; Zenobi, R. *J. Mass Spectrom.* **2008**, *43*, 185–195.

(5) Gobom, J.; Mueller, M.; Egelhofer, V.; Theiss, D.; Lehrach, H.; Nordhoff, E. *Anal. Chem.* **2002**, *74*, 3915–3923.

(6) Moskovets, E.; Chen, H.-S.; Pashkova, A.; Rejtar, T.; Andreev, V.; Karger, B. L. *Rapid Commun. Mass Spectrom.* **2003**, *17*, 2177–2787.

(7) Belgacem, O.; Buchacher, A.; Pock, K.; Josic, D.; Sutton, C.; Rizzi, A.; Allmaier, G. *J. Mass Spectrom.* **2002**, *37*, 1118–1130.

(8) Müller, R.; Allmaier, G. *Rapid Commun. Mass Spectrom.* **2006**, *20*, 3803–3806.

(9) Kemptner, J.; Marchetti-Deschmann, M.; Müller, R.; Ivens, A.; Turecek, P.; Schwarz, H. P.; Allmaier, G. *Rapid Commun. Mass Spectrom.* **2010**, *24*, 761–767.

(10) Kemptner, J.; Marchetti-Deschmann, M.; Müller, R.; Ivens, A.; Turecek, P.; Schwarz, H.-P.; Allmaier, G. *Recombinant van Willebrand factor as molecular weight marker for mass spectrometry analysis*, US 2010/0155588 A1. June 24. 2010.

(11) Ludwig, K.; Habbach, S.; Krieglstein, J.; Klumpp, S.; König, S. *PLoS ONE* **2011**, *6*, e23612.

(12) Greiner, M.; Carter, P.; Corn, B.; Zink, D. *Nat. Meth.* **2004**, *1*, 87–89.

(13) Müller, R.; Marchetti-Deschmann, M.; Elgass, H.; Breiteneder, H.; Kratzmeier, M.; Allmaier, G. *Electrophoresis* **2010**, *31*, 3850–3862.

(14) Goetz, H.; Kuschel, M.; Wulff, T.; Sauber, C.; Miller, C.; Fisher, S.; Woodward, C. *J. Biochem. Biophys. Methods* **2004**, *60*, 281–293.

(15) Müller, R.; Marchetti, M.; Kratzmeier, M.; Elgass, H.; Kuschel, M.; Zenker, A.; Allmaier, G. *Anal. Bioanal. Chem.* **2007**, *389*, 1859–1868.

(16) Seyfried, B. K.; Siekmann, J.; Turecek, P. L.; Schwarz, H. P.; Scheiflinger, F.; Zappe, H.; Bossard, M. L.; Allmaier, G. *Int. J. Mass Spectrom.* **2011**, *305*, 157–163.

(17) Farmer, T. B.; Caprioli, R. M. *J. Mass. Spectrom.* **1998**, *33*, 697–704.

(18) Helin, J.; Calentey, J.; Kalkkinen, N.; Bamford, D. H. *Rapid Commun. Mass Spectrom.* **1999**, *13*, 185–190.

(19) Chen, F.; Mädler, S.; Weidmann, S.; Zenobi, R. *J. Mass Spectrom.* **2012**, *47*, 560–566.

(20) Bich, C.; Bovet, C.; Rochel, N.; Peluso-Iltis, C.; Panagiotidis, A.; Nazabal, A.; Moras, D.; Zenobi, R. *J. Am. Soc. Mass Spectrom.* **2010**, *21*, 635–645.

(21) Seyfried, B. K.; Siekmann, J.; Belgacem, O.; Wenzel, R. J.; Turecek, P. L.; Allmaier, G. *J. Mass Spectrom.* **2010**, *45*, 612–617.

(22) Doyle, D. A.; Morais Cabral, J.; Pfuetzner, R. A.; Kuo, A.; Gulbis, J. M.; Cohen, S. L.; Chait, B. T.; MacKinnon, R. *Science* **1998**, *280*, 69–77.

(23) Lee, S.-Y.; Lee, A.; Chen, J.; MacKinnon, R. *Proc. Nat. Acad. Sci. USA* **2005**, *102*, 15441–15446.

(24) Dale, G. E.; Oefner, C.; D'Arcy, A. *J. Struct. Biol.* **2003**, *142*, 88–97.

(25) Malawski, G. A.; Hillig, R. C.; Monteclaro, F.; Eberspaecher, U.; Schmitz, A. A. P.; Crusius, K.; Huber, M.; Egner, U.; Donner, P.; Müller-Tiemann, B. *Protein Sci.* **2006**, *15*, 2718–2728.

(26) Forrer, P.; Binz, H. K.; Stumpp, M. T.; Plückthun, A. *ChemBioChem* **2004**, *5*, 183–189.

(27) Spurlino, J. C.; Lu, G.-Y.; Quiocho, F. A. *J. Biol. Chem.* **1991**, *266*, 5202–5219.

(28) Riggs, P. *Mol. Biotechnol.* **2000**, *15*, 51–63.

(29) Phillips, J. J.; Millership, C.; Main, E. R. *Angew. Chem. Int. Ed.* **2012**, *51*, 13132–13135.

(30) Binz, H. K.; Stumpp, M. T.; Forrer, P.; Amstutz, P.; Plückthun, A. *J. Mol. Biol.* **2003**, *332*, 489–503.

(31) Stumpp, M. T.; Binz, H. K.; Amstutz, P. *Drug Discov. Today* **2008**, *13*, 695–701.

(32) Oberhauser, A. F.; Marszalek, P. E.; Erickson, H. P.; Fernandez, J. M. **1998**, *393*, 181–185.

(33) Carrion-Vazquez, M.; Oberhauser, A. F.; Fowler, S. B.; Marszalek, P. E.; Broedel, S. E.; Clarke, J.; Fernandez, J. M. *Proc. Nat. Acad. Sci. USA* **1999**, *9*, 3694–3699.

(34) Carrion-Vazquez, M.; Li, H.; Li, H.; Marszalek, P. E.; Oberhauser, A. F.; Fernandez, J. M. *Nat. Struct. Biol.* **2003**, *10*, 738–743.

(35) Dietz, H.; Rief, M. *Proc. Nat. Acad. Sci. USA* **2004**, *101*, 16192–16197.

(36) Dietz, H.; Rief, M. *Proc. Nat. Acad. Sci. USA* **2006**, *103*, 244–1247.

(37) Arad-Haase, G.; Chuartzman, S. G.; Dagan, S.; Nevo, R.; Kouza, M.; Mai, B. K.; Nguyen, H. T.; Li, M. S.; Reich, Z. *Biophys. J.* **2010**, *99*, 238–247.

(38) Ingendoh, A.; Karas, M.; Hillenkamp, F.; Giessmann, U. *Int. J. Mass Spectrom. Ion Processes* **1994**, *131*, 345–354.

(39) Beavis, R. C.; Chait, B. T. *Anal. Chem.* **1990**, *62*, 1836–1840.

(40) Wenzel, R. J.; Röhling, U.; Nazabal, A.; Hillenkamp, F. *A Detector Device for High Mass Ion Detection, A Method for Analyzing Ions of High Mass and a Device for Selection Between Ion Detectors*, WO 2009/086642 A1. July 16. 2009.

(41) Winograd, N. *Anal. Chem.* **2005**, *77*, 142 A–149 A.

Supporting Information Available

This material is available free of charge via the Internet at `http://pubs.acs.org/`.

A New, Modular Mass Calibrant for High-Mass MALDI-MS

Simon Weidmann,[†] Konstantin Barylyuk,[†] Nadezhda Nespovitaya,[†] Stefanie Mädler,[†,‡] and Renato Zenobi[*,†]

Department of Chemistry and Applied Biosciences, ETH Zurich, CH-8093 Zurich, Switzerland

E-mail: zenobi@org.chem.ethz.ch

Supporting Information

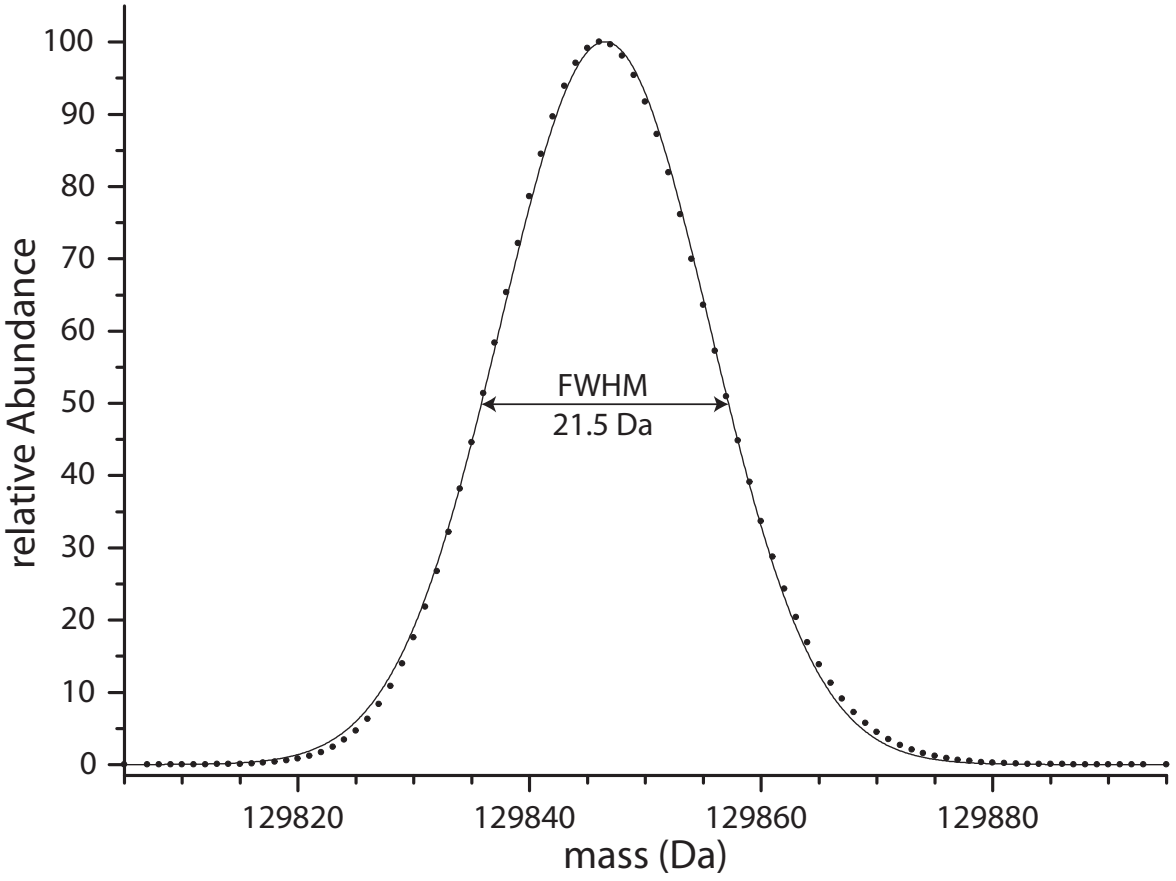
A schematic representation and the primary sequence of the newly developed modular high-mass calibrant MPB_3 is shown in Figure S 1.

Isotopic pattern distribution (Figure S 2) was calculated based the molecular formula of MBP_3 ($C_{5861}H_{9017}O_{1776}N_{1517}S_{22}$) using the algorithm provided by L. Patiny (EPFL, Lausanne, Switzerland) under `http://www.chemcalc.org`. The resolution setting "1" and the reference version "2012" were chosen for pattern simulation. To determine the FWHM a Gaussian peak was fitted using the algorithm provided by Origin 8.6 (OriginLab Corporation, Northampton MA, USA).

[*]To whom correspondence should be addressed
[†]Department of Chemistry and Applied Biosciences, ETH Zurich, CH-8093 Zurich, Switzerland
[‡]Current address: Centre for Research in Mass Spectrometry, York University, Toronto, ON M3J 2R7, Canada

1	MKIEEGKLVI	WINGDKGYNG	LAEVGKKFEK	DTGIKVTVEH	PDKLEEKFPQ	VAATGDGPDI	
61	IFWAHDRFGG	YAQSGLLAEI	TPDKAFQDKL	YPFTWDAVRY	NGKLIAYPIA	VEALSLIYNK	
121	DLLFNPPKTW	EEIPALDKEL	KAKGKSALMF	NLQEPYFTWP	LIAADGGYAF	KYENGKYDIK	
181	DVGVDNAGAK	AGLTFLVDLI	KNKHMNADTD	YSIAEAAFNK	GETAMTINGP	WAWSNIDTSK	
241	VNYGVTVLPT	FKGQPSKPFV	GVLSAGINAA	SPNKELAKEF	LENYLLTDEG	LEAVNKDKPL	
301	GAVALKSYEE	ELVKDPRIAA	TMENAQKGEI	MPNIPQMSAF	WYAVRTAVIN	AASGRQTVDE	
361	ALKDAQTNSS	SNNNNNNNNN	NLGDDDD/KVP	AAKVPMKIEE	GKLVIWINGD	KGYNGLAEVG	
421	KKFEKDTGIK	VTVEHPDKLE	EKFPQVAATG	DGPDIIFWAH	DRFGGYAQSG	LLAEITPDKA	
481	FQDKLYPFTW	DAVRYNGKLI	AYPIAVEALS	LIYNKDLLFN	PPKTWEEIPA	LDKELKAKGK	
541	SALMFNLQEP	YFTWPLIAAD	GGYAFKYENG	KYDIKDVGVD	NAGAKAGLTF	LVDLIKNKHM	
601	NADTDYSIAE	AAFNKGETAM	TINGPWAWSN	IDTSKVNYGV	TVLPTFKGQP	SKPFVGVLSA	
661	GINAASPNKE	LAKEFLENYL	LTDEGLEAVN	KDKPLGAVAL	KSYEEELVKD	PRIAATMENA	
721	QKGEIMPNIP	QMSAFWYAVR	TAVINAASGR	QTVDEALKDA	QTNSSSNNNN	NNNNNNLGGS	
781	GENLYFQ/GGS	GMKIEEGKLV	IWINGDKGYN	GLAEVGKKFE	KDTGIKVTVE	HPDKLEEKFP	
841	QVAATGDGPD	IIFWAHDRFG	GYAQSGLLAE	ITPDKAFQDK	LYPFTWDAVR	YNGKLIAYPI	
901	AVEALSLIYN	KDLLFNPPKT	WEEIPALDKE	LKAKGKSALM	FNLQEPYFTW	PLIAADGGYA	
961	FKYENGKYDI	KDVGVDNAGA	KAGLTFLVDL	IKNKHMNADT	DYSIAEAAFN	KGETAMTING	
1021	PWAWSNIDTS	KVNYGVTVLP	TFKGQPSKPF	VGVLSAGINA	ASPNKELAKE	FLENYLLTDE	
1081	GLEAVNKDKP	LGAVALKSYE	EELVKDPRIA	ATMENAQKGE	IMPNIPQMSA	FWYAVRTAVI	
1141	NAASGRQTVD	EALKDAQTNS	SSNNNNNNNN	NNLGCGSHHH	HHH		

S 1: In the primary sequence of the recombinantly expressed MBP_3 polyprotein, the three MBP subunits (blue background) are highlighted. The two different recognition sites are marked in green (ENTKrs) and red (TEVrs), respectively. On the C-terminus a single cysteine amino acid (black background) is inserted as well as a His-tag (orange) to allow for nickel affinity purification.



S 2: Isotopic distribution of MBP_3 calculated based on the molecular formula $C_{5861}H_{9017}O_{1776}N_{1517}S_{22}$. The relative abundance is calculated for every mass unit and Gaussian fit (line) is applied to these data points. The full width at half maximum is determined as 21.5 dalton.